



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB II

LANDASAN TEORI

2.1. *Cyberbullying*

Cyberbullying adalah penyalahgunaan *internet* untuk melecehkan, mengancam, mempermalukan, dan mengejek orang lain. (Wisnubrata, 2019) *Cyberbullying* adalah tindakan perundungan yang dilakukan pada dunia maya melalui barang elektronik seperti ponsel, komputer, dan *tablets*.

Cyberbullying bisa melalui SMS, teks, aplikasi, media sosial, forum, bahkan permainan *online* yang dimana orang lain dapat melihat, berpartisipasi, atau berbagi konten. *Cyberbullying* juga termasuk mengirimkan, memposting, atau membagikan konten negatif, berbahaya, palsu, atau tindakan jahat terhadap orang lain. Ini mencakup berbagi informasi pribadi yang dapat menyebabkan rasa malu atau penghinaan. (Stopbullying.gov, 2020)

Cyberbullying biasanya bukan hanya komunikasi satu kali, ini “terjadi secara berulang kali”, kecuali jika itu adalah sebuah ancaman pembunuhan atau ancaman serius terhadap keselamatan orang. Pada *cyberbullying* pelaku tidak bisa melihat respon langsung dari si korban sehingga dapat mengurangi kepuasan pelaku yang didapatkan dengan melihat sakit yang ditimbulkannya pada si korban, namun hal ini juga dapat mengurangi rasa empati dari pelaku terhadap si korban. (Rahayu, 2012)

Terdapat juga penelitian yang menyatakan bahwa 32 persen remaja mengatakan pernah melakukan *cyberbullying* dengan alasan mengisengi temannya dan media yang paling banyak digunakan adalah situs media sosial. (Rahayu, 2012)

2.2. *Twitter*

Twitter adalah situ berita dan jejaring sosial online tempat orang berkomunikasi dalam pesan singkat yang disebut *tweet*. *Tweeting* memposting pesan singkat untuk siapa saja yang mengikuti anda di *twitter*, dengan harapan kata – kata anda bermanfaat dan menarik bagi seseorang di *audiens* anda. Deskripsi lain tentang *twitter* dan *tweeting* mungkin *microblogging*. (Gil, 2020)

Twitter didirikan dan diresmikan pada tahun 2006 tepatnya pada tanggal 21 Maret 2006. Didirikan oleh Jack Dorsey, *twitter* sudah cukup di kenal di dunia, bahkan pada tahun 2014 *twitter* sudah menjadi salah satu dari 5 besar situs yang paling sering dikunjungi oleh banyak orang. (Basri, 2017)

Twitter sendiri memiliki beberapa fitur, seperti *tweet* atau kicauan yang merupakan fitur utama di *twitter* untuk mengirim dan melihat kicauan setiap pengguna *twitter*. *Following* merupakan fitur untuk mengikuti pengguna lain pada *twitter*. *Followers*, fitur yang digunakan untuk melihat siapa yang mengikuti anda pada media sosial *twitter*. *Biography* merupakan fitur yang digunakan untuk mengetahui pesan akun *twitter* anda pada *profile*. *Profile* merupakan salah satu fitur utama pada *twitter*, fitur ini digunakan untuk melihat *avatar twitter*, *biography twitter*, dan lainnya. Pada *twitter* juga terdapat fitur *top trending* yang digunakan untuk mempermudah penggunaanya untuk melihat *tweet* yang paling populer dan paling sering di-*tweet* oleh para pengguna *twitter* lainnya. (Basri, 2017)

2.3. *Data Mining*

Data Mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. (Muchlisin Riadi, 2017)

Data Mining memiliki dua fungsi utama, yaitu fungsi *descriptive* dan fungsi *predictive*. Fungsi *descriptive* adalah fungsi untuk memahami lebih jauh tentang data yang diamati. Dengan melakukan sebuah proses diharapkan bisa mengetahui perilaku dari sebuah data tersebut. (romadhoni, 2020) Sedangkan fungsi *predictive* yaitu sebuah proses yang nantinya digunakan untuk menemukan pola tertentu pada suatu data, lalu pola tersebut dapat diketahui dari berbagai *variable – variable* yang ada pada data. (romadhoni, 2020)

Kelebihan data mining, biasanya di mulai dari proses yang cukup panjang dan rumit, dari proses tersebut biasanya menghasilkan sesuatu yang baru, yang biasanya tidak diketahui sebelumnya. Sesuatu yang baru ini akan menambah pengetahuan para pengguna ataupun penelitian dan tentunya akan bermanfaat karena dapat digunakan untuk melakukan tindakan tertentu. Membuat analisa suatu data besar menjadi lebih mudah, sehingga mempermudah pengambilan keputusan atau prediksi data di masa depan. (Bootupacademyai, 2019)

2.4. Text Mining

Text Mining adalah salah satu proses yang termasuk dalam kateogi “*data science*” *text mining* atau yang lebih dikenal dengan istilah *text analytics* adalah teknologi *Artificial Intelligence* atau AI yang memungkinkan penggunaanya untuk mengubah konten inti dari sebuah dokumen teks menjadi sebua data kuantitatif secara cepat. (Wijaya, 2018)

Tujuan dari *text mining* adalah untuk mendapatkan informasi yang berguna dari sekumpulan dokumen. Jadi, sumber data yang digunakan pada *text mining* adalah sekumpulan teks yang memiliki *format* yang tidak terstruktur atau minimal semi terstruktur. Adapun tugas khusus dari *text mining* antara lain yaitu pengkategorisasian teks (*text categorization*) dan pengelompokan teks (*text clustering*). (Nindito, 2016)

Salah satu langkah pertama dalam proses *text mining* adalah mengatur dan menyusun data dengan cara tertentu sehingga dapat menjadi sasaran analisis kualitatif dan kuantitatif. Melakukannya secara khusus melibatkan penggunaan teknologi *Natural Language Processing* (NLP), yang menerapkan prinsip – prinsip *linguistic* komputasional untuk mengurai dan menginterpretasikan set data. (Arni, 2018)

Text Mining memiliki beberapa tipe antara lain (Abbott, 2013) :

2.4.1. Search and Information Retrieval

Information retrieval adalah ilmu mencari informasi yang relevan terhadap kebutuhan pengguna dari suatu kumpulan informasi secara otomatis. Contoh aplikasi umum yang digunakan dari *information retrieval* adalah *search engine* maupun mesin pencari yang terdapat pada jaringan internet. (Markijar, 2015)

2.4.2. Document Clustering

Document clustering adalah proses pengelompokan dokumen yang memiliki kesamaan topik. Tujuan dari proses *clustering* ini membagi dokumen berdasarkan kesamaan, sehingga memudahkan dalam proses pencarian. (Informatikalogi, 2017)

2.4.3. Document Classification

Document classification adalah bentuk tindakan penandaan dokumen untuk dimasukkan ke dalam kategori sesuai dengan kontennya. *Document classification* bisa dilakukan secara manual maupun otomatis, dan digunakan untuk mempermudah menyortir dan mengelola teks, gambar, atau video. (Pascual, 2019)

2.4.4. Web Mining

Web mining adalah ekstraksi pola – pola penting dan bermanfaat namun tersimpan secara implisit pada kumpulan data yang relatif besar pada

layanan *world wide web*. *Web mining* terdiri dari tiga bagian, yaitu : *web content mining*, *web structure mining*, dan *web usage mining*. (Nindito, 2016)

Web content mining adalah suatu proses otomatis untuk menemukan informasi yang berguna dari dokumen atau data. Pada intinya proses ini mengekstraksi kata kunci yang terkandung pada dokumen, ada 2 teknik yang biasanya di gunakan, pertama langsung melakukan *mining* terhadap data, dan kedua melakukan pencarian serta melakukan *improvisation* hasil pencarian seperti layaknya *search engine*. (Nindito, 2016)

Web structure mining dikenal juga sebagai *web log mining* adalah teknik yang digunakan untuk menemukan struktur *link* dari *hyperlink* dan membangun rangkuman *website* dan halaman *web*. (Nindito, 2016)

Kemudian *web usage mining* adalah teknik yang digunakan untuk mengenali perilaku pelanggan dan struktur *web* melalui informasi yang diperoleh dari *log*, *click stream*, *cookies*, dan *query*. Berbagai tool yang sudah ada antara lain *WebLogMiner* yang digunakan untuk *mining* data *log*. (Nindito, 2016)

2.4.5. Information Extraction

Information Extraction merupakan proses identifikasi fragmen – fragmen khusus suatu dokumen yang membentuk inti dari isi semantiknya. *Information Extraction* atau disingkat IE biasanya bergantung pada aturan ekstraksi yang dibuat khusus untuk sumber informasi tertentu, yang biasa disebut *wrapper*, *wrapper* digunakan untuk membentuk struktur sumber data agar dapat diproses. (Nurdiyanto, 2013)

Dalam praktiknya, *Information Extraction* sering dibantu dengan penggunaan pola tata bahasa serta aturan linguistik yang diterapkan pada bahasa alami manusia. Namun dengan bantuan domain pengetahuan yang sesuai konteks, representasi pengetahuan yang dihasilkan oleh proses IE dapat berkurang dari sisi kompleksitasnya. Selain itu, algoritma

pembelajarannya serta efisiensi ekstraksinya menjadi semakin baik. (Nurdiyanto, 2013)

2.4.6. *Natural Language Processing*

Natural Language Processing atau NLP merupakan kecerdasan buatan dan bahasa yang berkaitan dengan interaksi antara komputer dan bahasa alami manusia, dengan teknologi yang ada pada *natural language processing*, komputer dapat memahami bahasa manusia, demikian juga sebaliknya, komputer dapat membuat bahasa yang dimengerti manusia. (Nurdiyanto, 2013)

Fungsi dasar dalam *natural language processing* adalah *parsing*. *Parsing* atau analisis sintak merupakan proses analisis simbol baik pada basa alami maupun bahasa komputer, sesuai dengan aturan bahasa formal. (Nurdiyanto, 2013)

2.4.7. *Concept Extraction*

Concept Extraction adalah proses untuk mengidentifikasi frasa yang merujuk pada konsep minat dalam teks yang tidak terstruktur. *Concept extraction* merupakan salah satu komponen penting dalam melakukan proses teks otomatis. (Manabu Torii, 2011)

2.5. *Text Preprocessing*

Tahap *text preprocessing* atau tahap praproses data merupakan proses untuk mempersiapkan data mentah sebelum dilakukan proses lainnya. Pada umumnya *text preprocessing* data dilakukan dengan cara mengeliminasi data yang tidak sesuai atau mengubah data menjadi bentuk yang lebih mudah diproses oleh sistem. (Mujilahwati, 2016)

Tidak ada aturan pasti tentang setiap tahapan di dalam proses *text preprocessing* semua tergantung dari jenis data dan hasil yang diinginkan. Namun pada umumnya tahapan pada proses *text preprocessing* adalah *case*

folding, tokenization, filtering, stopword removal, dan stemming (Yulio, 2019)

2.5.1. Case Folding

Case Folding digunakan untuk mengkonversi keseluruhan teks dalam dokumen menjadi suatu bentuk standar (huruf kecil atau *lowercase*). (Informatikalogi, 2017)

2.5.2. Tokenizing

Tokenizing atau tokenisasi merupakan suatu proses pemotongan *string* atau kata pada suatu kalimat dan semua tanda baca dan tanda hubung akan dihilangkan. Proses ini bertujuan untuk memisahkan tiap kata agar dapat membedakan karakter – karakter tertentu yang diperlukan sebagai pemisah kata atau bukan. Proses tokenisasi mengandalkan karakter spasi pada dokumen sebagai pemisah kata. (Informatikalogi, 2017)

2.5.3. Filtering

Tahap *filtering* merupakan tahap pengambilan kata – kata penting dari hasil token. Bisa digunakan algoritma *stoplist* atau *wordlist*. *Stoplist* atau *wordlist* merupakan kata – kata yang tidak deskriptif yang dapat dibuang dalam pendekatan *bag of word*. (Informatikalogi, 2017)

2.5.4. Stemming

Teknik *stemming* digunakan untuk mengetahui akar sebuah kata. Teknik *stemming* diperlukan untuk memperkecil jumlah indeks yang berbeda dari suatu dokumen, juga untuk mengelompokkan kata – kata lain yang memiliki kata dasar dan arti yang serupa namun memiliki bentuk yang berbeda karena memiliki imbuhan yang berbeda. (Informatikalogi, 2017)

2.5.5. *Term Weighting*

Term weighting adalah prosedur yang terjadi selama proses pengindeksan teks untuk menilai setiap istilah untuk dokumen. *Term weighting* adalah penugasan nilai numerik pada istilah yang mewakili kepentingannya dalam dokumen untuk meningkatkan efektivitas pengambilan data. (El-Khair, 2009)

2.6. Teori Pengolahan Data

2.6.1. *Information Gain*

Information Gain adalah sebuah metode yang digunakan untuk menghitung nilai *entropy* yang ada dengan cara membagi *dataset* berdasarkan nilai dari *random variable*. (Brownlee, 2019) *Information Gain* juga memiliki keuntungan, jika informasi yang dimiliki semakin banyak maka akan didapatkan nilai *entropy* yang semakin rendah dan meminimalisir kejutan.

Pada teori informasi dapat di gambarkan bahwa “*surprise*” yang semakin tinggi dalam suatu kejadian. Probabilitas lebih rendah untuk terjadinya kejadian “*surprise*” dengan cara memperbanyak jumlah informasi atau memiliki informasi yang lebih banyak, sedangkan distribusi probabilitas terjadi dimana peristiwa dan kemungkinan lebih “*surprising*” dan “*entropy*” yang lebih besar. (Astuti, 2017)

Pemilihan fitur dengan *information gain* dilakukan dalam tiga tahapan, pertama menghitung nilai *information gain* untuk setiap atribut dalam dataset original. Kemudian menentukan batas atau *threshold* yang diinginkan. Hal ini dilakukan agar memungkinkan atribut yang berbobot sama dengan batas atau lebih besar akan dipertahankan serta membuang atribut yang berada di bawah batas. Lalu memperbaiki dataset dengan mengurangi atribut yang ada. (Muchammad Rifqi Maulana, 2015)

2.6.2. Support Vector Machine

Support Vector Machine (SVM) pertama kali diperkenalkan oleh Vapnik pada tahun 1992 sebagai rangkaian konsep – konsep unggulan dalam bidang *pattern recognition*. Sebagai salah satu metode yang di bilang muda. SVM sudah memiliki kemampuan sebagai *state of the art* dalam *pattern recognition*. SVM sendiri merupakan metode *learning machine* yang berkerja atas prinsip *Structural Risk Minimization* (SRM) dengan tujuan menemukan *hyperplane* terbaik yang memisahkan dua *class* pada *input space*. (Anto Satriyo Nugroho, 2003)

2.6.2.1. Karakteristik Support Vector Machine (SVM)

Secara prinsip *Support Vector Machine* merupakan *linear classifier*, lalu SVM memiliki *pattern recognition* yang dilakukan dengan mentransformasikan data pada *input space* ke ruang yang memiliki dimensi yang lebih tinggi. SVM juga menerapkan strategi *Structural Risk Minimization*, SVM juga memiliki prinsip kerja yang hanya mampu menangani klasifikasi dua kelas. (Anto Satriyo Nugroho, 2003)

2.6.2.2. Kelebihan dan Kekurangan Support Vector Machine

1. Generalisasi

Generalisasi di definisikan sebagai kemampuan suatu metode untuk melakukan klasifikasi suatu *pattern*, yang tidak termasuk data yang di pakai dalam fase pembelajaran metode tersebut. (Anto Satriyo Nugroho, 2003)

2. Curse of Dimensionality

Curse of dimensionality di definisikan sebagai masalah yang dihadapi suatu metode *pattern recognition* dalam melakukan estimasi parameter dikarena jumlah sampel yang relative sedikit dibandingkan dimensional ruang vector data tersebut. (Anto Satriyo Nugroho, 2003)

Curse of dimensionality sering dialami dalam aplikasi di bidang *bioimmedical engineering*, karena biasanya data biologi yang terjadi sangat terbatas, dan penyediaannya memerlukan biaya tinggi, sehingga SVM digunakan untuk membuktikan bahwa tingkat generalisasi yang diperoleh SVM tidak di pengaruhi oleh dimensi dari *input vector*. (Anto Satriyo Nugroho, 2003)

3. Feasibility

Support Vector Machine dapat di implementasikan relative mudah, dikarenakan proses penentuan *support vector* bias dirumuskan dalam QP *problem*, dengan demikian jika kita memiliki *library* untuk menyelesaikan QP *problem*, dengan sendirinya SVM dapat diimplementasikan dengan mudah. Selain itu dapat diselesaikan dengan metode sekuensial. (Anto Satriyo Nugroho, 2003)

4. Landasan teori

Sebagai metode yang berbasis statistik, SVM memiliki landasan teori yang dapat dianalisa dengan jelas, dan tidak bersifat tertutup. (Anto Satriyo Nugroho, 2003)

2.7. Tools

2.7.1. Software R

Software R adalah suatu *software* untuk menganalisis data statistic, yang hamper sama seperti SPSS, MINITAB, S-Plus, SAS, atau Eviews. R termasuk kelompok *software statistic open source* yang tidak memerlukan lisensi, atau lebih dikenal dengan *freeware*. (IGSD, 2012)

Salah satu keunggulan R adalah komunitas besar yang tergabung dalam satu *mailing-list*, dokumentasi para pengguna yang mudah di akses, grup *stack overflow* yang sangat aktif, dan *packages R* yang dibagikan oleh sesame pengguna. *Software R* juga sangat baik digunakan dalam visualisasi data. (IYKRA, 2018)

2.7.2. Microsoft Excel

Microsoft excel adalah sebuah program atau aplikasi yang merupakan dari paket instalasi *Microsoft Office*, berfungsi untuk mengolah angka menggunakan *spreadsheet* yang terdiri dari baris dan kolom untuk mengeksekusi perintah. *Microsoft Excel* telah menjadi *software* pengolah data atau angka terbaik di dunia, selain itu *Microsoft Excel* telah didistribusikan secara multi-platform. *Microsoft excel* tidak hanya tersedia di *platform Windows*, *Microsoft Excel* juga tersedia di *Mac OS*, dan *Android*. (Mitrakuliah.com, 2019)

2.8. Penelitian Sebelumnya

No	Nama	Judul Jurnal	Volume	Edisi	Tahun	Nama Jurnal	Metode	Hasil Penelitian	Kesimpulan
1	(Astuti, 2017)	Seleksi Atribut Menggunakan <i>Information Gain</i> untuk <i>Clustering</i> Penduduk Miskin Dengan <i>Validity Index Xie Beni</i>	Volume 6	1	2017	Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer.	<i>Information Gain</i>	Hasil pengujian terhadap 23,500,100, dan 1313 untuk jumlah <i>cluster</i> 2,3,4,5,6, dan 7 menunjukkan bahwa nilai dari <i>Xie-Beni Index</i> terkecil adalah 5 dengan nilai 0.1343, sehingga <i>cluster</i> yang paling optimal	Penggunaan <i>Information Gain</i> dengan <i>threshold</i> 0.0001 untuk <i>Clustering</i> dengan menghilangkan atribut penghasil memiliki hasil <i>cluster</i> yang sama dengan menggunakan atribut penghasil.

No	Nama	Judul Jurnal	Volume	Edisi	Tahun	Nama Jurnal	Metode	Hasil Penelitian	Kesimpulan
								adalah 5.	
2	(Zahro Malihah, 2018)	Perilaku <i>cyberbullying</i> pada remaja dan kaitannya dengan control diri dan komunikasi orang tua	Volume 11	2	2018	Jur. Ilm. Keluarga & Konsumen.	<i>Cross Sectional Study</i>	Hasil penelitian menunjukkan bahwa komunikasi orang tua – remaja dan kontrol diri remaja berhubungan negatif dengan perilaku <i>cyberbullying</i> remaja, sehingga diharapkan orang tua lebih memperhatikan cara dengan intensitas berkomunikasi dengan remaja.	Komunikasi orang tua dengan remaja dan control diri berpengaruh negatif dengan <i>cyberbullying</i> tetapi komunikasi dengan anak dan orang tua diharapkan dapat menyalurkan nilai positif sehingga anak terhindar dari perilaku menyimpang.
3	(Billy Kaleb Hananto, 2018)	Penerapan <i>Maximum TF-IDF Normalization</i> Terhadap Metode KNN Untuk Klasifikasi	Volume 2	12	2018	Jurnal Pengembangan Teknologi Informasi dan Komputer	TF-IDF, KNN	Penelitian ini membuktikan bahwa dengan perubahan yang dilakukan pada metode Max TF-IDF tidak memberik	Penelitian ini mencoba menerapkan metode bernama <i>Maximum TF-IDF</i> , yang bekerja untuk membagi nilai frekuensi dengan jumlah kata

No	Nama	Judul Jurnal	Volume	Edisi	Tahun	Nama Jurnal	Metode	Hasil Penelitian	Kesimpulan
		asi Dataset <i>Multiclass Panichella</i> Pada <i>Review Aplikasi Mobile</i>						an hasil yang memuaskan, dikarenakan tidak meningkatkan nilai akurasi dari KNN <i>Cosine Similarity</i> terhadap dataset <i>Panichella</i> .	terbanyak yang muncul, dengan harapan bahwa metode KNN dapat memberikan hasil akurasi yang lebih tinggi dengan di ubahnya metode TF-IDF.
4	(Syafitri Hidayatul Annur Aini, 2018)	Seleksi Fitur <i>Information Gain</i> untuk Klarifikasi Penyakit Jantung Menggunakan Kombinasi Metode <i>K-Nearest Neighbor</i> dan <i>Naive Bayes</i>	Volume 2	9	2018	Jurnal Pengembangan teknologi Informasi dan Ilmu Komputer	<i>Information Gain, KNN, Naive Bayes</i>	Hasil penelitian tentang klasifikasi penyakit jantung menggunakan seleksi fitur <i>Information Gain</i> dengan kombinasi KNN dan <i>Naive Bayes</i> , yaitu nilai akurasi tertinggi yang diperoleh saat menggunakan data latih dengan tabel kelas seimbang,	Kesimpulan yang dapat diambil bahwa algoritma <i>Information Gain</i> dengan kombinasi <i>K-Nearest Neighbor</i> dan <i>Naive Bayes</i> dapat digunakan untuk klasifikasi penyakit jantung.

No	Nama	Judul Jurnal	Volume	Edisi	Tahun	Nama Jurnal	Metode	Hasil Penelitian	Kesimpulan
								<p>yaitu sebesar 92.31% saat menggunakan enam fitur dengan nilai $K=25$. Enam fitur yang digunakan antara lain, <i>thal</i>, jenis nyeri dada, <i>flourosopy</i>, rata – rata detak jantung, <i>oldpeak</i> dan <i>exercise included angina</i>. Sedangkan nilai akurasi tertinggi yang diperoleh saat menggunakan data latih dengan label kelas tidak seimbang, yaitu sebesar 92.31% saat mengguna</p>	

No	Nama	Judul Jurnal	Volume	Edisi	Tahun	Nama Jurnal	Metode	Hasil Penelitian	Kesimpulan
								kan empat fitur dengan nilai K=35	
5	(Agnes Rossi Trisna Lestari, 2017)	Analisis Sentimen Tentang Opini Pilkada DKI 2017 Pada Dokumen Twitter Berbahasa Indonesia menggunakan <i>Naive Bayes</i> dan Pembobotan <i>Emoji</i>	Volume 1	12	2017	Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer	<i>Naive Bayes</i>	Pengujian akurasi pada penelitian untuk pembobotan tekstual sebesar 68.52% dan pada pembobotan non-tekstual sebesar 75.93% dan pada penggabungan kedua pembobotan menjadi 74.81% sehingga disimpulkan bahwa penggabungan kedua pembobotan meningkatkan hasil akurasi sistem, dari 11 kali pengujian di dapat hasil penambahan akurasi pada	Kesimpulan yang dapat diambil bahwa algoritma <i>naive bayes</i> dan pembobotan <i>emoji</i> dapat diterapkan pada penelitian tersebut. Pada data latih dan data uji dilakukan deteksi <i>emoji</i> dimana ditemukan maka dilakukan pembobotan <i>emoji</i> dan normalisasi min-max, kemudian praproses teks, dan pembobotan tekstual, sehingga di dapatkan hasil berupa sentimen positif maupun negatif.

No	Nama	Judul Jurnal	Volume	Edisi	Tahun	Nama Jurnal	Metode	Hasil Penelitian	Kesimpulan
								pembandingan konstanta penggali tertentu.	